



# Combinatorial trees arising in the study of interval exchange transformations

Julien Cassaigne, Sébastien Ferenczi, Luca Q. Zamboni

## ► To cite this version:

Julien Cassaigne, Sébastien Ferenczi, Luca Q. Zamboni. Combinatorial trees arising in the study of interval exchange transformations. European Journal of Combinatorics, 2011, 32, 10.1016/j.ejc.2011.05.002 . hal-01265525

**HAL Id: hal-01265525**

**<https://hal.science/hal-01265525>**

Submitted on 3 Feb 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# COMBINATORIAL TREES ARISING IN THE STUDY OF INTERVAL EXCHANGE TRANSFORMATIONS

JULIEN CASSAIGNE, SÉBASTIEN FERENCZI, AND LUCA Q. ZAMBONI

**ABSTRACT.** In this paper we study a new class of combinatorial objects we call *trees of relations* equipped with an operation we call *induction*. These trees were first introduced in [3] in the context of interval exchange transformations but they may be studied independently from a purely combinatorial point of view. They possess a variety of interesting combinatorial properties and have already been linked to a number of different areas including ergodic theory and number theory [3, 4]. In a recent sequel to this paper, R. Marsh and S. Schroll have established interesting connections to the theory of cluster algebras and polygonal triangulations [5]. For each tree of relations  $G$ , we let  $\Gamma(G)$  denote the smallest set of trees of relations containing  $G$  and invariant under induction. The induction mapping allows us to endow  $\Gamma(G)$  with the structure of a connected directed graph, which we call the graph of graphs. We investigate the structure of  $\Gamma(G)$  and define a circular order based on the tree structure which turns out to be a complete invariant for the induction mapping. This gives a complete characterization of  $\Gamma(G)$  which allows us to compute its cardinality in terms of Catalan numbers. We show that the circular order also defines an abstract secondary structure similar to one occurring in genetics in the study of RNA.

## 1. INTRODUCTION

In [3] we introduced a new induction algorithm for a family of interval exchange transformations  $T$  in the hyperelliptic Rauzy class. This algorithm, called the *self-dual induction*, provides new insight on the symbolic dynamics of the trajectories [4]. Our aim was to describe completely the trajectories of points, and to relate both the combinatorial and dynamical properties of the underlying system to the number-theoretic properties of an associated multi-dimensional continued fraction algorithm. If  $T$  is an exchange on  $k$  intervals, then at each stage of our induction, we induce  $T$  (by first return) on a disjoint union of  $k - 1$  sub-intervals  $E_j$ , each containing the point  $\beta_j$  of discontinuity of  $T^{-1}$  and whose endpoints are in the orbits of the discontinuities of  $T$ . This process defines a multi-dimensional continued fraction algorithm generated by the  $2k - 2$  parameters  $\{l_j, r_j\}_{1 \leq j \leq k-1}$  where  $l_j$  is the distance from  $\beta_j$  to the left endpoint of  $E_j$  and  $r_j$  the distance from  $\beta_j$  to the right endpoint of  $E_j$ .

As soon as  $k \geq 3$ , the  $2k - 2$  parameters  $\{l_j, r_j\}_{1 \leq j \leq k-1}$  are not independent and in fact satisfy  $k - 2$  symmetric relations of the form  $l_i = l_j$ , or  $r_i = r_j$  or  $l_i + r_i = l_j + r_j$  for some  $i \neq j$ . At each state of the induction, these relations, which are a consequence of the isometry of the transformation  $T$  and the nature of the underlying permutation, may be coded by a tree on  $k - 1$  nodes (labelled 1 through  $k - 1$ ) with labelled edges, where the labels take on three possible values corresponding to the three different types of relations. Thus a labelled edge in the tree between nodes  $i$  and  $j$  indicates a relation between parameter  $x \in \{l_i, r_i\}$  and  $y \in \{l_j, r_j\}$ , and the exact

---

*Date:* April 26, 2011.

*2000 Mathematics Subject Classification.* Primary 68R10; Secondaries 05C05, 37E05.

*Key words and phrases.* Interval exchange transformations, induction, trees of relations, Catalan numbers, RNA secondary structures, pseudoknot-free.

form of the relation is given by the edge label. These trees, we call *trees of relations*, are at the very core of the dynamics of hyperelliptic interval exchange transformations, and in fact in [3] we show that the entire combinatorial description of the trajectories of  $T$  may be deduced directly from them.

In the present paper we define and study trees of relations, equipped with the operation of induction, from a purely combinatorial view, that is removed from the context of interval exchange transformations. Very simply, a tree of relations is a tree in which each edge is labelled by either  $+$ ,  $=$ , or  $-$ , and such that no two adjacent edges have the same label. Figure 1 depicts an example of a tree of relations with ten vertices. These trees, equipped with the operation of induction, constitute a new discrete structure possessing rich combinatorial properties. Together they define directed graphs whose vertices consist of trees of relations with vertices 1 through  $k - 1$ , and where the directed edges between vertices are given by the induction mapping.

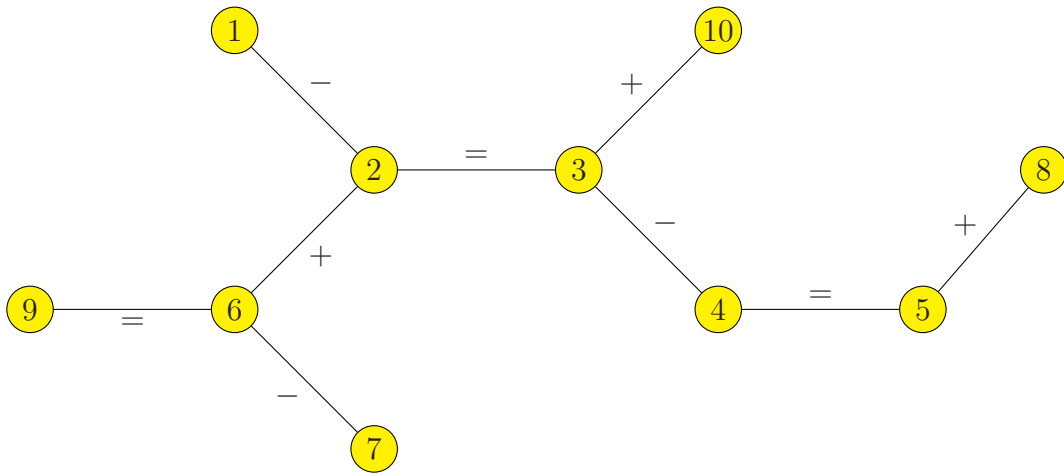


FIGURE 1. A tree of relations on 10 vertices.

An outline of the paper is as follows: In §2 we consider an example of an interval exchange  $T$  on 4-intervals, to illustrate the induction algorithm in the context of interval exchanges.

In §3 we define and study the basic properties of trees of relations. We define the induction mapping in purely combinatorial terms as a mapping from trees of relations to trees of relations.

In §4 we show that for every tree of relations  $G$ , the set  $\Gamma(G)$ , defined as the smallest set of trees of relations containing  $G$  and invariant under induction, may be endowed with the structure of a connected directed graph. We call the directed graph  $\Gamma(G)$  the graph of graphs of  $G$ .

In §5, §6 and §7, we investigate the structure of the graph of graphs  $\Gamma(G)$ . For this purpose we introduce in §5 two auxiliary notions: *shapes* and *fillings*: A shape is a tree of relations in which the vertices are unlabelled, they represent the skeleton of the tree, while the filling represents the passage from shapes to trees. We then show that the trees in  $\Gamma(G)$  realise every possible shape.

In §6 we define a circular order on the vertices of a tree of relations which is determined by its tree structure. We show that two fillings of a shape are in the same  $\Gamma(G)$  if and only if they define the same circular order on their respective vertices: the circular order is a complete invariant for the induction mapping, and thus gives a full characterization of  $\Gamma(G)$ .

In §7 we use this invariant to count both the number of shapes as well as the cardinality of  $\Gamma(G)$ , by formulas involving Catalan numbers.

In §8 we discuss an interesting connection between the circular structure defined in §6 and a similar structure occurring in genetics in the study of RNA.

Recently, R. Marsh and S. Schroll have written a sequel to this paper further extending the combinatorial theory developed herein and establishing interesting and surprising connections to polygonal  $m$ -angulations, Fuss-Catalan combinatorics, and the theory of cluster algebras (see [5]).

## 2. INTERVAL EXCHANGE TRANSFORMATIONS

Let us consider the interval exchange transformation  $T$  on 4-intervals as shown in Figure 2 (by convention, all intervals are open on the right, closed on the left). The transformation  $T$  maps by isometry the first interval  $[0, 1 - \beta_3[$  onto the interval  $[\beta_3, 1[$ , the second interval  $[1 - \beta_3, 1 - \beta_2[$  onto  $[\beta_2, \beta_3[$ , the third interval  $[1 - \beta_2, 1 - \beta_1[$  onto  $[\beta_1, \beta_2[$ , and the fourth interval  $[1 - \beta_1, 1[$  onto  $[0, \beta_1[$ .

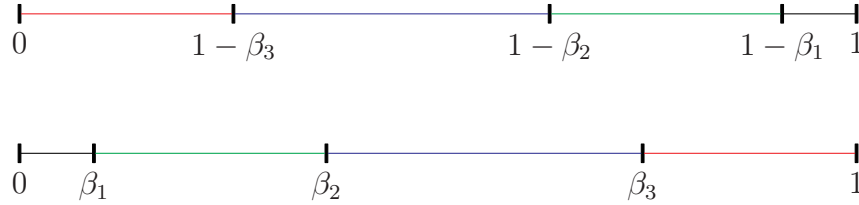


FIGURE 2. A symmetric 4-interval exchange transformation.

For convenience, we further impose the initial condition

$$0 < \beta_1 < 1 - \beta_3 < \beta_2 < 1 - \beta_2 < \beta_3 < 1 - \beta_1$$

so that  $\beta_1$  is in the interval  $E_1 = [0, 1 - \beta_3[$ ,  $\beta_2$  in the interval  $E_2 = [1 - \beta_3, 1 - \beta_2[$ , and  $\beta_3$  in the interval  $E_3 = [1 - \beta_2, 1 - \beta_1[$ . For each  $j \in \{1, 2, 3\}$  we consider the two parameters  $l_j, r_j$  where  $l_j$  and  $r_j$  are defined as the respective distances between the point  $\beta_j$  and the left and right endpoints of  $E_j$ ; thus  $|E_j| = l_j + r_j$  is the length of  $E_j$ .

We remark that there are two relations between these parameters, namely that  $r_1 = r_3$  and  $l_2 = l_3$ ; they are a consequence of the underlying isometry of  $T$  and the choice of the permutation by which we re-arrange the intervals. We record (or code) them as follows: for  $r_1 = r_3$  we write  $1 \hat{-} 3$  (or equivalently  $3 \hat{-} 1$ ) and for  $l_2 = l_3$  we write  $3 \hat{+} 2$  (or equivalently  $2 \hat{+} 3$ ). We may combine these two expressions by forming a tree with vertices  $\{1, 2, 3\}$  and with an undirected edge labelled  $-$  between 1 and 3 and one labelled  $+$  between 2 and 3 as shown in Figure 3; this tree is denoted also by  $1 \hat{-} 3 \hat{+} 2$  or  $2 \hat{+} 3 \hat{-} 1$  (see the beginning of §3).

The self-dual induction defined in [3] starts from the three intervals  $E_{1,0} = E_1$ ,  $E_{2,0} = E_2$ ,  $E_{3,0} = E_3$ , and creates three smaller intervals  $E_{1,1}$ ,  $E_{2,1}$ ,  $E_{3,1}$  (thus they are no longer adjacent). By iteration, we obtain three families of nested intervals  $E_{1,n}$ ,  $E_{2,n}$ ,  $E_{3,n}$ . At each step of the induction we consider the sub-interval  $E_{i,n}$  containing the special point  $\beta_i$ , and recalculate the corresponding parameters  $l_i, r_i$ . It turns out that at each stage there will be two relations between the parameters of the following form: for some  $i \neq j$ , either  $l_i = l_j$ , which we code by  $i \hat{+} j$  or equivalently by  $j \hat{+} i$ , or  $r_i = r_j$ , which we code by  $i \hat{-} j$  or equivalently by  $j \hat{-} i$ , or  $|E_i| = |E_j|$ , or equivalently  $l_i + r_i = l_j + r_j$ , which we code by  $i \hat{=} j$  or equivalently by  $j \hat{=} i$ .

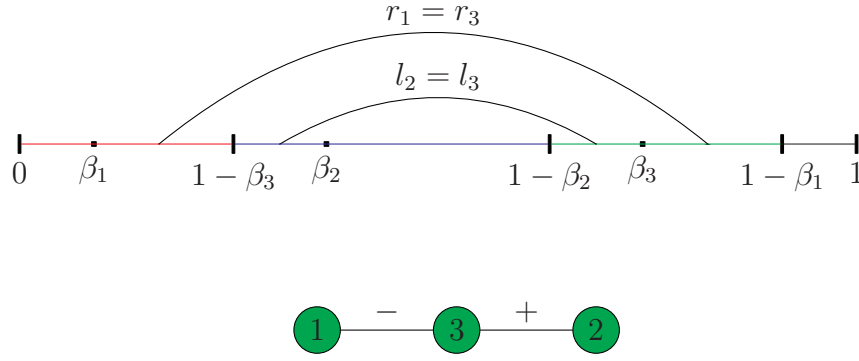


FIGURE 3. The coding of the parameters in state 0.

The complete definition of the self-dual induction was made in [3]; as it is not necessary to the understanding of the present paper, we choose to follow first what happens on one example and postpone the full definition to §3 below, as it requires the full definition of the induction operation on trees in §2 below. Thus, suppose that in state 0 (the initial state) we have  $r_1 > l_1$ ,  $r_2 > l_2$ ,  $r_3 > l_3$ . Then, in passing to the subsequent state (state 1), each interval  $E_i$  is cut from the right by the amount  $l_i$ , as shown in Figure 4. It follows from the previous relations that the new parameters satisfy the new relations  $l_2 = l_3$  (or  $2 \hat{+} 3$ ) and  $|E_1| = |E_3|$  (or  $1 \hat{=} 3$ ).

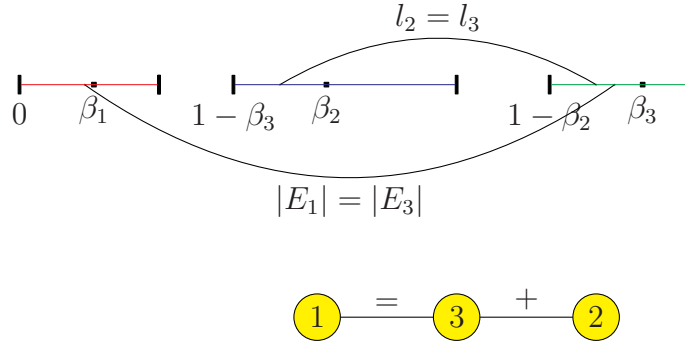


FIGURE 4. The coding of the parameters in state 1.

Suppose that in state 1, the corresponding parameters satisfy  $l_1 > r_3$ ,  $r_2 > l_2$ ,  $l_3 > r_1$  (this happens whenever the parameters in state 0 satisfy  $l_1 + l_3 > r_1 = r_3 > \max(l_1, l_3)$  and  $r_2 > 2l_2$ , which can be realised). Then the interval  $E_2$  is cut from the right by the amount  $l_2$ , while the other intervals are not cut. Although the new parameter values of  $l_2$  and  $r_2$  differ from the corresponding values in the previous state, these two parameters were not involved in the preceding relations and hence the coding remains unchanged as shown in Figure 5.

Suppose that in state 2 the corresponding parameters satisfy  $l_1 > r_3$ ,  $l_2 > r_2$ ,  $l_3 > r_1$  (again there are initial values of the parameters for which this happens). To go to state 3, the interval  $E_1$  is cut from the left by the amount  $r_3$ , and  $E_3$  is cut from the left by  $r_1$ , while  $E_2$  is cut from the left

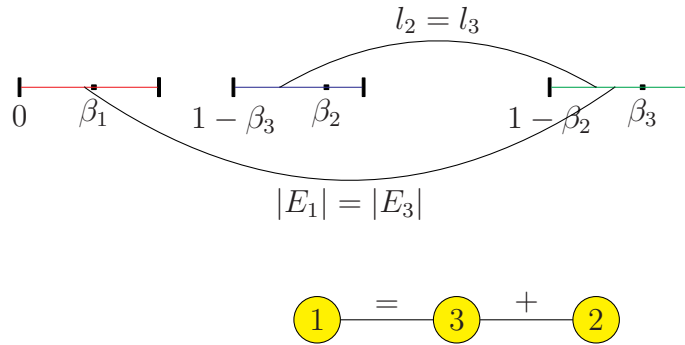


FIGURE 5. The coding of the parameters in state 2.

by  $r_2$ . This gives rise to the coding  $3\hat{+}1\hat{=2}$  as shown in Figure 6.

Suppose that in state 3 the corresponding parameters satisfy  $l_1 < r_2, l_2 < r_1, l_3 < r_3$ . In passing from Figure 6 to Figure 7 the interval  $E_1$  is cut from the right by the amount  $l_2$  and  $E_2$  is cut from the right by  $l_1$ , while  $E_3$  is cut from the right by  $l_3$ .

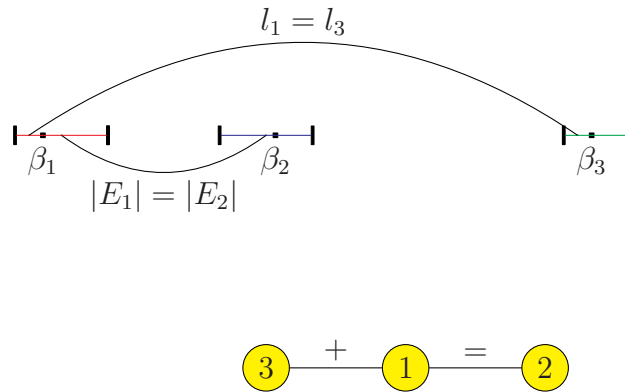


FIGURE 6. The coding of the parameters in state 3.

As it turns out, each state has been coded by a tree of relations. We next build a graph whose vertices consist of all trees of relations coding the possible states, and where there is a labelled directed edge between any two adjacent (or consecutive) states. The edges are labelled by either  $+$  if the intervals are cut from the left, or  $-$  if they are cut from the right. The resulting graph is shown in Figure 17 below, where the highlighted edges represent the initial path between consecutive

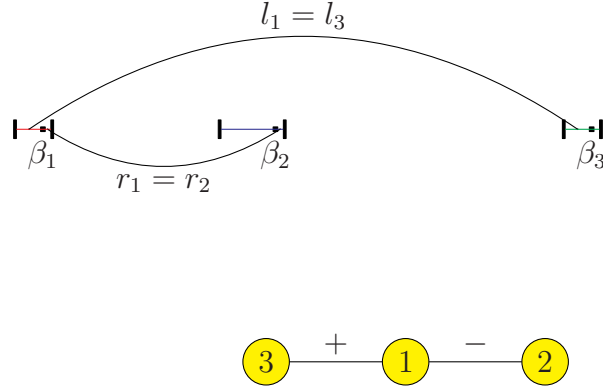


FIGURE 7. The coding of the parameters in state 4.

states outlined in the present example. If we iterate the self-dual induction infinitely many times we obtain an infinite path in this graph.

### 3. TREES OF RELATIONS AND INDUCTION

By a *tree* we mean a non-oriented connected graph which has no cycles.

**Definition 3.1.** A *tree of relations* on a finite nonempty set  $K$  is a tree  $G$  satisfying the following three conditions:

- The vertices of  $G$  are the elements of  $K$ .
- Each edge of  $G$  is labelled with  $\{+, =, -\}$ .
- No two adjacent edges of  $G$  have the same label.

**Notations.** Throughout this paper, we consider edges labelled with  $\{+, =, -\}$ . We use the notation  $a \hat{+} b$ , resp.  $a \hat{=} b$ ,  $a \hat{-} b$ , to denote the edge labelled  $+$ , resp.  $=$ ,  $-$ , between the vertices  $a$  and  $b$ . By further abbreviation, in describing a given tree of relations  $G$  we write just (for example) that  $a \hat{+} b$  in  $G$  to express that there is an edge  $a \hat{+} b$  in  $G$ , and  $a \hat{+} b \hat{=} c$  instead of  $a \hat{+} b$  and  $b \hat{=} c$ . The hats are used only to avoid writing expressions like  $1 = 2$  or  $1 - 2 = 3$ , and thus are not needed in pictures or in expressions like  $a + \text{edge}$ . Clearly  $a \hat{+} b$  is equivalent to  $b \hat{+} a$ , and the same if we replace  $+$  by  $-$  or  $=$ .

**Example 3.2.** The tree given in Figure 1 can be described in many equivalent ways, for example  $1 \hat{=} 2 \hat{=} 3 \hat{+} 10$ ,  $2 \hat{+} 6 \hat{=} 9$ ,  $6 \hat{-} 7$ ,  $3 \hat{-} 4 \hat{=} 5 \hat{+} 8$ , or alternatively  $8 \hat{+} 5 \hat{=} 4 \hat{-} 3 \hat{=} 2 \hat{+} 6 \hat{-} 7$ ,  $6 \hat{=} 9$ ,  $2 \hat{-} 1$ ,  $10 \hat{+} 3$ .

To each tree of relations  $G$  on  $K$  we associate three bijections  $s, t, u : K \rightarrow K$  defined as follows: for each  $a \in K$  we put

- $s(a) = b$  if  $a \hat{=} b$  for some  $b \neq a$ . Otherwise set  $s(a) = a$ .
- $t(a) = b$  if  $a \hat{+} b$  for some  $b \neq a$ . Otherwise set  $t(a) = a$ .
- $u(a) = b$  if  $a \hat{-} b$  for some  $b \neq a$ . Otherwise set  $u(a) = a$ .

It is immediate that

$$s^2 = t^2 = u^2 = Id.$$

**Definition 3.3.** Let  $G$  be a tree of relations. By a *positive chain* we mean a maximal connected subtree of  $G$  having no  $-$  edges. Similarly a *negative chain* is a maximal connected subtree of  $G$  having no  $+$  edges. By a *signed chain* we mean either a positive or a negative chain.

Figures 8 and 9 illustrate the positive and negative chains respectively of the tree of relations given in Figure 1.

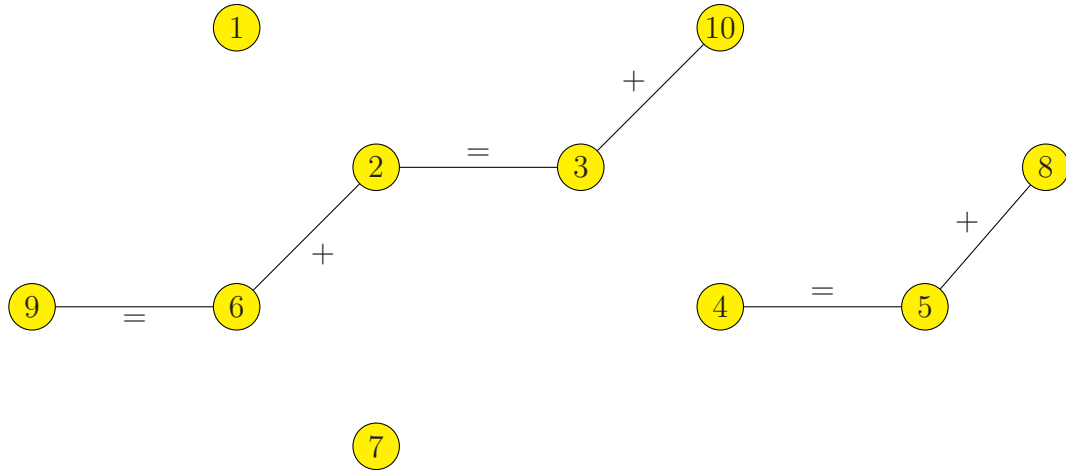


FIGURE 8. The positive chains in Figure 1.

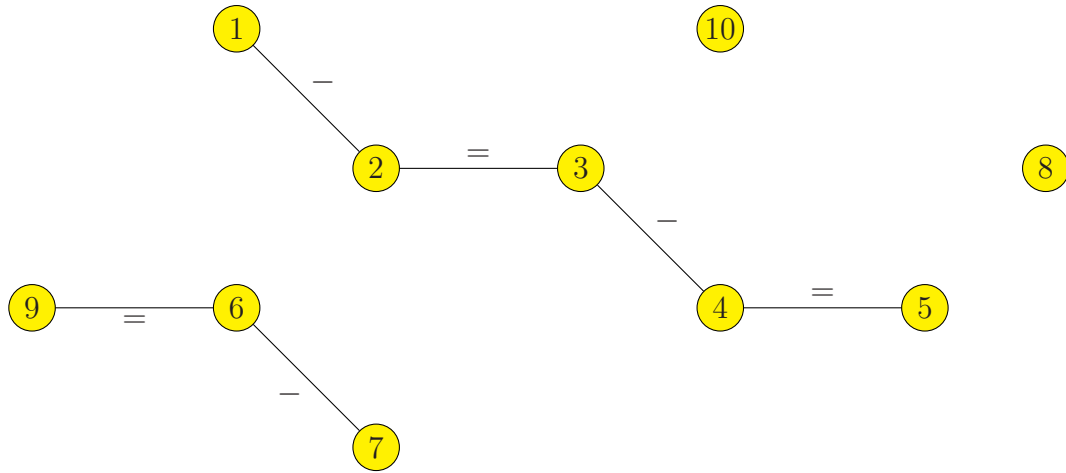


FIGURE 9. The negative chains in Figure 1.

We next define an operation on trees of relations which we call *induction*. This operation associates to each tree of relations  $G$  on  $K$  and each signed chain  $B$  of  $G$ , a tree of relations  $J_B(G)$  on  $K$  as follows:

**Definition 3.4.** Let  $G$  be a tree of relations on  $K$ , with first bijection  $s$ , and  $B$  be a signed chain. The tree of relations  $J_B(G)$  is defined by the vertices  $a$ ,  $a \in K$ , and the following edges



- if  $a \in B$ ,  $b \in B$ , and  $a \hat{+} b$  in  $G$ , then  $s(a) \hat{=} s(b)$  in  $J_B(G)$ ,
- if  $a \in B$ ,  $b \in B$ , and  $a \hat{-} b$  in  $G$ , then  $s(a) \hat{=} s(b)$  in  $J_B(G)$ ,
- if  $a \in B$ ,  $b \in B$ ,  $B$  is a positive chain, and  $a \hat{=} b$  in  $G$ , then  $a \hat{+} b$  in  $J_B(G)$ ,
- if  $a \in B$ ,  $b \in B$ ,  $B$  is a negative chain, and  $a \hat{=} b$  in  $G$ , then  $a \hat{-} b$  in  $J_B(G)$ ,
- if  $a \notin B$  or  $b \notin B$ , and  $a \hat{R} b$  in  $G$ , then  $a \hat{R} b$  in  $J_B(G)$ , for  $R \in \{+, =, -\}$ .

It is readily verified that  $J_B(G)$  is a tree of relations. We note that  $B' = J_B(B)$  is again a signed chain in  $J_B(G)$ . The mapping  $(G, B) \mapsto J_B(G)$  may be described geometrically in three simple steps as shown in Figures 11-13: We consider the tree of relations in Figure 1 and the highlighted negative chain  $B$  consisting of vertices  $\{1, 2, 3, 4, 5\}$  (see Figure 10). The resulting tree is shown in Figure 14.

The induction on a disjoint union of signed chains, though it is not effectively used in the present paper, is an important tool in the definition of the induction on interval exchange transformations (§3 below):

**Definition 3.5.** For a union  $B = B_1 \dots \cup B_k$  of piecewise disjoint signed chains, we define  $J_B(G)$  as the composition  $J_{B_k} \circ \dots \circ J_{B_1}(G)$ , which by Definition 3.4 above is independent of the order of the  $B_i$ .

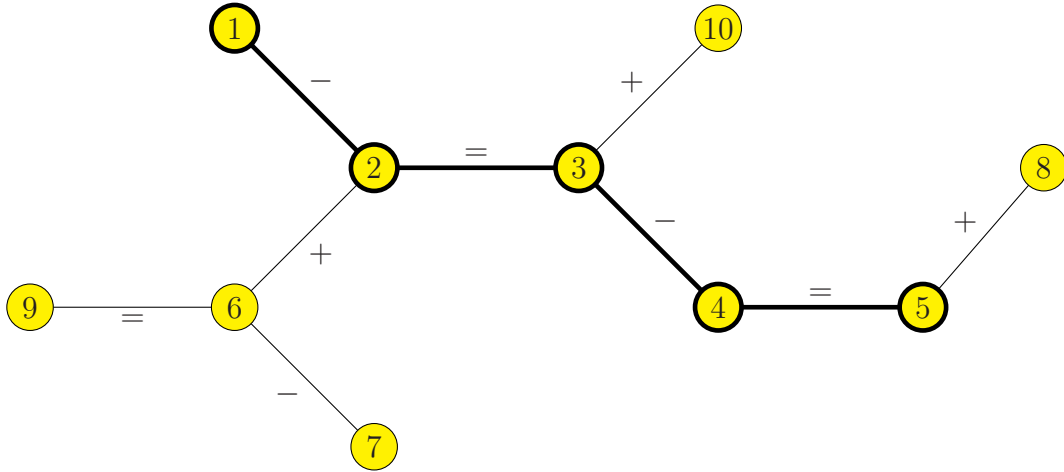
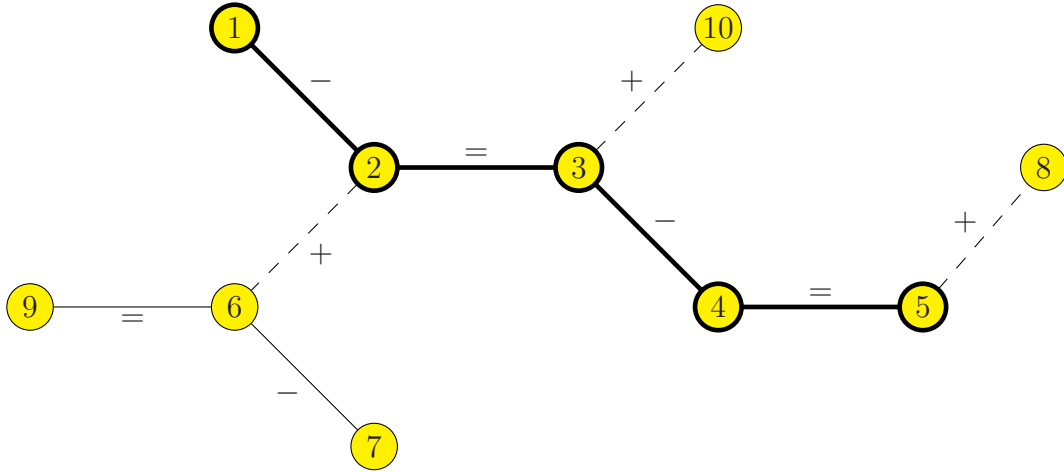
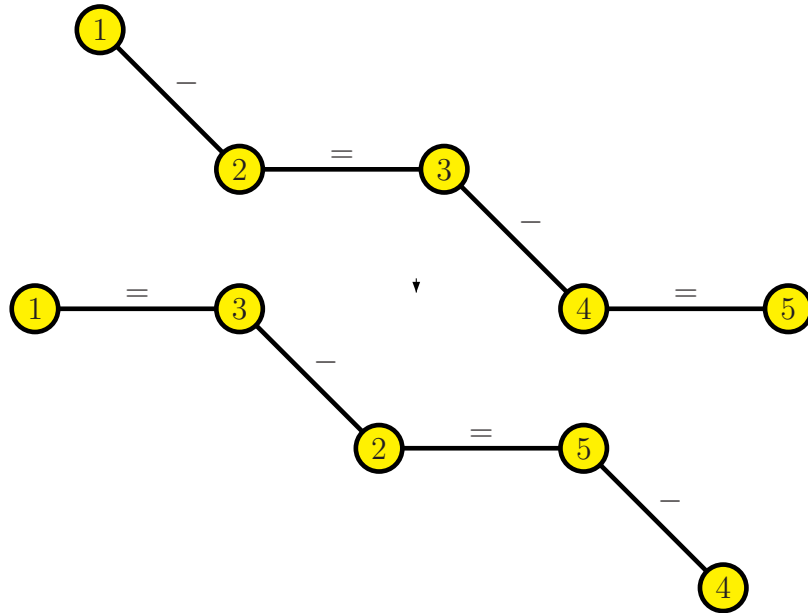


FIGURE 10. An induction on the tree of Figure 1.

FIGURE 11. Step 1: Prune auxiliary branches to isolate  $B$ .FIGURE 12. Step 2: Exchange vertices  $a \leftrightarrow s(a)$  and edges  $- \leftrightarrow =$ .

**Example 3.6.** We look again at the trees in §2 above.

Let  $G$  be  $1\hat{-}3\hat{+}2$ . Then  $s = (123)$ ,  $t = (132)$ ,  $u = (321)$ . The positive chains are  $B_1 = 3\hat{+}2$ ,  $B_2 = 1$ , the negative chains are  $B_3 = 1\hat{-}3$  and  $B_4 = 2$ .  $J_{B_1}(G)$  is  $1\hat{-}3\hat{=}$ ,  $J_{B_3}(G)$  is  $1\hat{=}3\hat{+}2$ ,  $J_{B_2}(G)$  and  $J_{B_4}(G)$  are  $G$ . When we went from state 0 to 1, we induced successively on the two negative chains  $B_3$  and  $B_4$ , thus by Definition 3.5 on the union  $B_3 \cup B_4$ .

Let  $G$  be  $1\hat{=}3\hat{+}2$ . Then  $s = (321)$ ,  $t = (132)$ ,  $u = (123)$ . The only positive chain is  $B_1 = 1\hat{=}3\hat{+}2$ , the negative chains are  $B_2 = 1\hat{-}3$  and  $B_3 = 2$ .  $J_{B_1}(G)$  is  $2\hat{=}1\hat{+}3$ ,  $J_{B_2}(G)$  is  $1\hat{-}3\hat{+}2$ ,  $J_{B_3}(G)$  is  $G$ . When we went from state 1 to 2, we induced on the negative chain  $B_3$ , and from state 2 to state 3 we induced on the positive chain  $B_1$ .

Let  $G$  be  $2\hat{=}1\hat{+}3$ ; to go from state 3 to 4, we induced on the union of the two negative chains  $2\hat{-}1$  and 3.

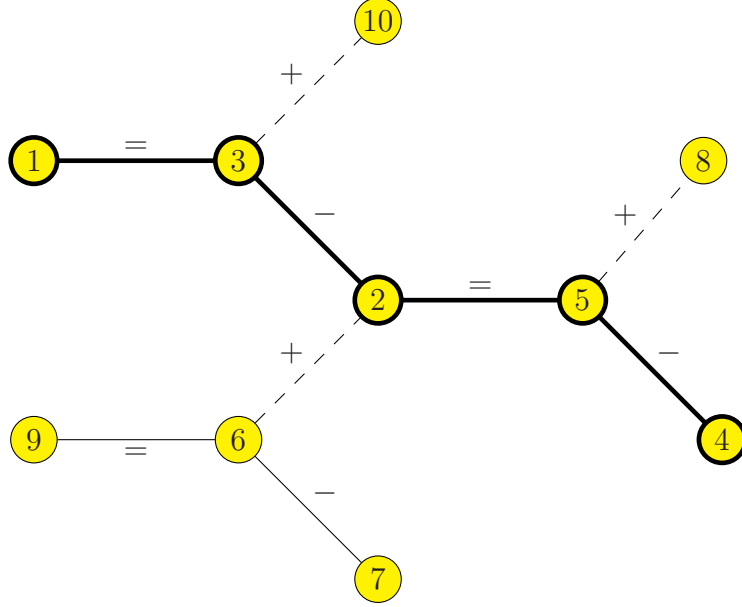


FIGURE 13. Step 3: Re-join pruned branches to their original connections.

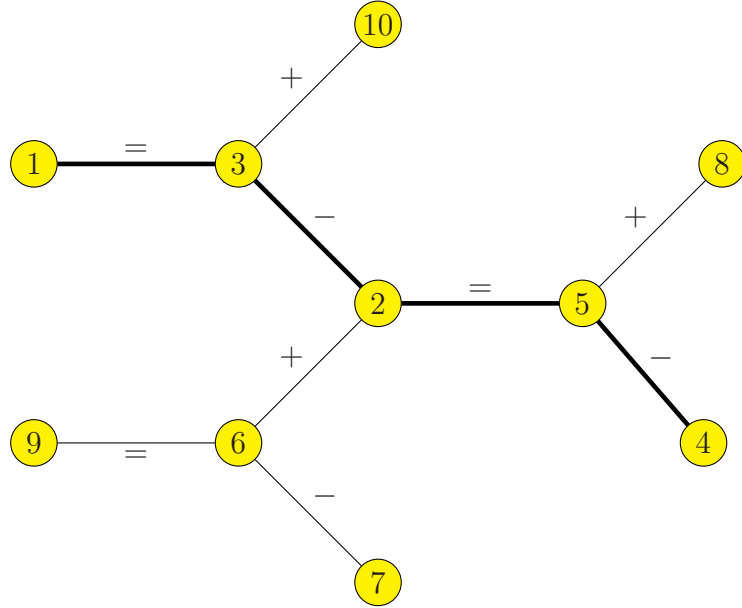


FIGURE 14. The resulting tree of relations.

We can now give the full rules of the self-dual induction in [3]. At a given stage, they depend both on the relations between  $l_i$  and  $r_i$ , coded by a tree of relations  $G$  with first bijection  $s$ , and on the signs of the  $l_i - r_{s(i)} = l_{s(i)} - r_i$ . Namely, let  $B$  be the union of all positive chains on which  $l_i > r_{s(i)}$  on every vertex, and of all negative chains on which  $l_i < r_{s(i)}$  on every vertex; then, if  $i$  is a vertex of a positive chain in  $B$ ,  $E_i$  is cut on the left by an amount  $r_{s(i)}$ , while if  $i$  is a vertex of a negative chain in  $B$ ,  $E_i$  is cut on the right by an amount  $l_{s(i)}$ ; if  $i$  is a vertex outside of  $B$ ,  $E_i$  is

not cut. It is proved in [3] that *the tree of relations at the next stage is  $J_B(G)$* . Note that the two other bijections of the tree  $G$  are also used in [3], but for technical reasons there we use slightly different maps, namely  $p = ts$  and  $m = us$ .

As is explained in [3], these rules, and the definitions of trees and induction, come naturally from the requirements of the self-dual induction, and this theory needs the whole machinery of trees of relations to work satisfactorily.

The proof of the following lemma is immediate from the above definitions:

**Lemma 3.7.** *Let  $B$  a signed chain of a tree of relations  $G$ . The bijections  $(s', t', u')$  of  $J_B(G)$  are given as follows:*

- if  $a \in B$  and  $B$  is a positive chain,,  $s'(a) = sts(a)$ ,  $t'(a) = s(a)$ ,  $u'(a) = u(a)$ ;
- if  $a \in B$  and  $B$  is a negative chain,,  $s'(a) = sus(a)$ ,  $t'(a) = t(a)$ ,  $u'(a) = s(a)$ ;
- if  $a \notin B$ ,  $s'(a) = s(a)$ ,  $t'(a) = t(a)$ ,  $u'(a) = u(a)$ .

#### 4. GRAPHS OF GRAPHS

In this section we use the induction mapping to construct a graph whose vertices consist of trees of relations with  $k$  vertices, and where the directed edges between vertices are defined in terms of the induction mapping.

**Definition 4.1.** *For a given tree of relations  $G$ , let  $\Gamma(G)$  be the smallest set of trees of relations on  $K$  which contains  $G$  and is closed under induction.*

We give  $\Gamma(G)$  the structure of an oriented graph as follows: for each  $G' \in \Gamma(G)$  and each signed edge  $B$  of  $G'$ , we place a directed edge from  $G'$  to  $J_B(G')$  labelled  $+$  (resp.  $-$ ) if  $B$  is a positive (resp. negative) chain. We call  $\Gamma(G)$  a *graph of graphs*.

Let  $k$  be a positive integer. We define the *initial tree of relations* on the set  $K = \{1, 2, \dots, k\}$ , denoted  $G_0(k)$ , by

$$G_0(k) = 1 \hat{-} k \hat{+} 2 \hat{-} (k-1) \hat{+} 3 \hat{-} (k-2) \dots$$

As we saw in §2 for  $k = 3$ , this is the tree of relations which codes the relations between the parameters of the self-dual induction in the initial state for an interval exchange transformation on  $k+1$  intervals, with suitable conditions on the permutation which re-arranges the intervals, and on the positions of the discontinuities.

We denote by  $\Gamma(k)$  the graph of graphs  $\Gamma(G_0(k))$ . Figures 15, 16 and 17 illustrate  $\Gamma(1)$ ,  $\Gamma(2)$  and  $\Gamma(3)$  (the highlighted edges in  $\Gamma(3)$  correspond to the example in §2). We shall see later that  $\Gamma(4)$  has 28 vertices while  $\Gamma(5)$  has 90 vertices.

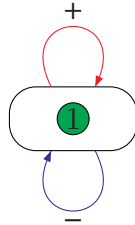
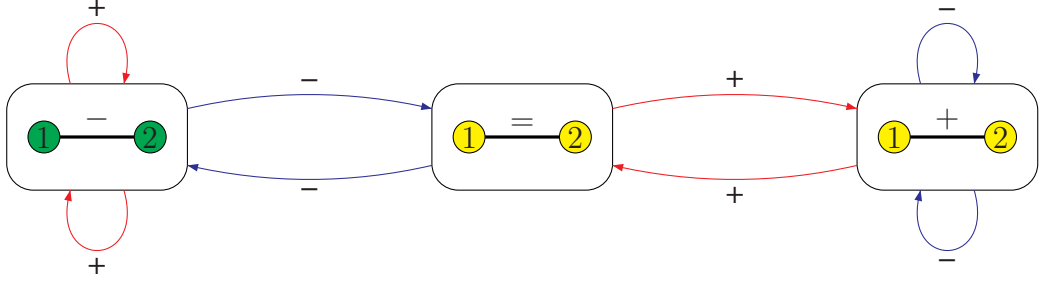
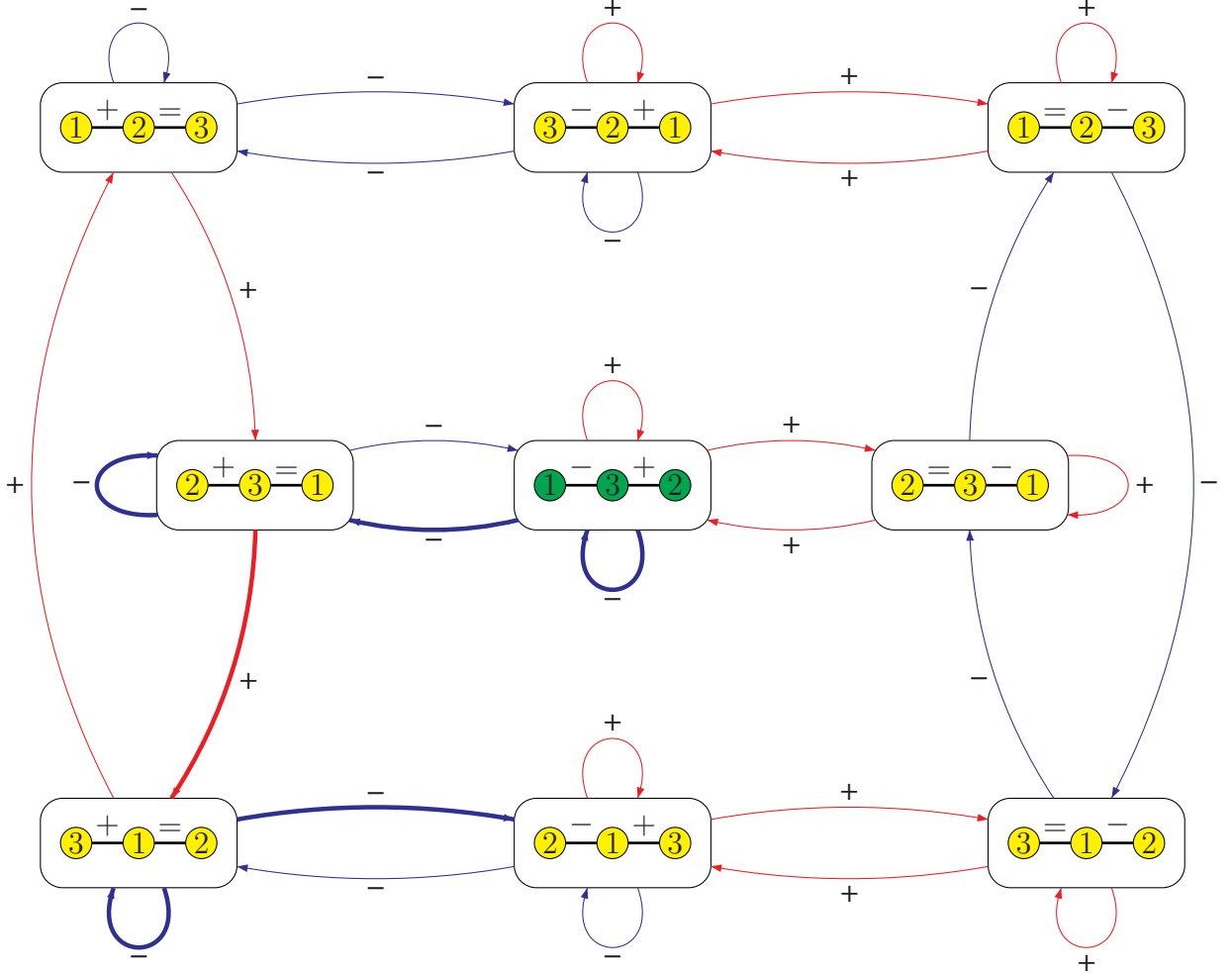


FIGURE 15. The graph of graphs  $\Gamma(1)$ .

FIGURE 16. The graph of graphs  $\Gamma(2)$ .FIGURE 17. The graph of graphs  $\Gamma(3)$ .

## 5. SHAPES & FILLINGS

In the next three sections we shall investigate the structure of  $\Gamma(k)$ , and in particular determine its cardinality. We begin by introducing the following two auxiliary notions.

**Definition 5.1.** A shape is a (non-oriented) tree  $F$  with  $k$  unnamed vertices and  $k-1$  edges labelled  $+$ ,  $-$ , or  $=$ , such that two adjacent edges never have the same label.

The shape of a tree of relations  $G$  is the tree obtained from  $G$  by removing the labels of the vertices. We say that  $G$  is a filling of its shape.

A rooted tree of relations (resp. shape) is a tree of relation (resp. shape) together with the choice of one vertex, called the root.

**Proposition 5.2.** For any tree of relations  $G$ , there exists a tree of relations  $G_\star$  without any  $=$  edge and a sequence of inductions  $J_{B_1}, \dots, J_{B_n}$  such that  $G_\star = J_{B_n} \dots J_{B_1}(G)$ .

**Proof**

If  $G$  has no  $=$  edge, we are done; otherwise, we shall relate, by a sequence of inductions,  $G$  to a tree of relations  $G'$  having one less  $=$  edge.

An  $=$  edge between two vertices  $a$  and  $b$  has at most four adjacent edges  $a \hat{-} y_1, a \hat{+} y_2, b \hat{-} y_3, b \hat{+} y_4$ ; for  $y \in \{y_1, \dots, y_4\}$  let  $G_y(a, b)$  be the connected component containing  $y$  of the tree  $G$  deprived of the edge from  $y$  to its adjacent  $a$  or  $b$  vertex. We say that  $a \hat{=} b$  is an *extremal*  $=$  edge if there is at most one  $i$  such that  $G_{y_i}(a, b)$  has at least one  $=$  edge; if it exists, we call this  $G_{y_i}(a, b)$  the *dangerous* subtree of  $a \hat{=} b$ ; note that the non-dangerous  $G_{y_i}(a, b)$  contain no vertex with more than two adjacent edges, because three edges adjacent to one common vertex must have three different labels. For example, the  $=$  edges between vertices 9 and 6 and between vertices 4 and 5 in Figure 1 are extremal, while the  $=$  edge between vertices 2 and 3 is not extremal.

There is at least one extremal  $=$  edge, since otherwise we can follow a directed infinite path in  $G$  and  $G$  is not a finite tree.

Suppose  $a \hat{=} b$  is an extremal edge of  $G = G_0$ , and suppose first that it has a dangerous subtree and the edge leading to it is a  $+$  edge. We denote by  $b_0$  the vertex which is on that edge and on the extremal edge, and by  $a_0$  the other vertex on the extremal edge.

We build a finite sequence of graphs  $G_n$  such that

- $G_{n+1}$  is obtained from  $G_n$  by a sequence of one or two inductions of sign  $(-1)^{n+1}$ ;
- if  $n$  is even,  $G_n$  has an extremal  $=$  edge  $a_n \hat{=} b_n$ , with at most four adjacent edges  $a_n \hat{-} x_n, a_n \hat{+} y_n, b_n \hat{-} z_n, b_n \hat{+} t_n$ , such that  $t_n$  exists,  $G_{t_n}(a_n, b_n)$  is made of the union of  $G_{t_0}(a_0, b_0)$  and a branch from  $t_0$  to  $t_n$  with  $n$  vertices ( $t_0$  excluded,  $t_n$  included), and is the dangerous subtree of  $a_n \hat{=} b_n$ ;
- if  $n$  is odd, the same is true with all signs changed to the opposite.

This is true for  $n = 0$ , for a unique choice of  $(x_0, y_0, z_0, t_0)$ . Suppose  $n$  is even; then we consider the four possibilities for the negative chain containing the edge  $a_n \hat{=} b_n$ :

- if it is  $x_n \hat{-} a_n \hat{=} b_n \hat{-} z_n$ , we induce on it, getting the negative chain  $z_n \hat{-} a_n \hat{-} b_n \hat{-} x_n$ , and induce on this new chain, getting  $G_{n+1}$  with the negative chain  $a_n \hat{-} z_n \hat{=} x_n \hat{-} b_n$ . Then we put  $a_{n+1} = y_n, b_{n+1} = x_n$ .
- If it is  $x_n \hat{-} a_n \hat{=} b_n$ , we induce on it, getting the negative chain  $x_n \hat{-} b_n \hat{-} a_n$ , and induce on this new chain, getting  $G_{n+1}$  with the negative chain  $b_n \hat{-} x_n \hat{=} a_n$ . Then we put  $a_{n+1} = a_n, b_{n+1} = x_n$ .
- If it is  $a_n \hat{=} b_n \hat{-} z_n$ , we induce on it, getting  $G_{n+1}$  with the negative chain  $b_n \hat{-} a_n \hat{=} z_n$ . Then we put  $a_{n+1} = z_n, b_{n+1} = b_n$ .
- If it is  $a_n \hat{=} b_n$ , we induce on it, getting  $G_{n+1}$  with the negative chain  $b_n \hat{-} a_n$  and we stop the process.

For  $n$  odd we do the same construction, changing each sign to its opposite; and the sequence  $G_n$  has the claimed properties; note that  $G_{n+1}$  has the same number of  $=$  edges as  $G_n$ , except at the

last step where this number decreases by one. If the dangerous subtree of  $a \hat{=} b$  in  $G_0$  is linked to  $b_0$  by a  $-$  edge, we do the same process with all signs changed. If there is no dangerous subtree of  $a \hat{=} b$  in  $G_0$ , but  $G_0$  is not reduced to  $a \hat{=} b$ , we choose  $b_0$  to be  $a$  or  $b$ , and  $t_0$  such that  $b_0 \hat{e}_0 t_0$ , where  $e_0 = +$  or  $e_0 = -$ , and do the same construction as above, defining  $t_n$  by  $b_n \hat{e}_n t_n$  where  $e_n$  is the sign of  $(-1)^{n e_0}$ ; all our assertions remain true except that  $G_{t_n}(a_n, b_n)$  is not the dangerous subtree of  $a_n \hat{=} b_n$ . If  $G_0$  is reduced to  $a \hat{=} b$ , we define  $G_1$  as the tree  $a \hat{=} b$  and stop the process. In all cases, as all the  $G_n$  have the same finite number of vertices, the process has to stop, and the final  $G_n = G'$  has one  $=$  edge less than  $G_0$ .  $\square$

**Corollary 5.3.** *If  $G$  has  $k$  vertices, every possible shape with  $k$  vertices appears as the shape of a tree of relations in  $\Gamma(G)$ .*

**Proof**

By Proposition 5.2, it is enough to show that every shape without  $=$  edges appears. If  $k$  is odd, there is only one such shape, the shape of the initial tree  $\hat{-}.\hat{+}...\hat{+}$  and the result is proved.

If  $k$  is even, the shapes without  $=$  edges are  $\hat{-}.\hat{+}...\hat{-}$  and  $\hat{+}.\hat{-}...\hat{+}$ ; we have shown that  $\Gamma(G)$  contains one tree with one of these shapes, for example  $a_1 \hat{-} a_2 \hat{+} a_3 \dots a_{2p-1} \hat{-} a_{2p}$ ; then by a negative induction  $\Gamma(G)$  contains also  $a_1 \hat{+} a_2 \hat{-} a_3 \hat{-} \dots a_{2p-1} \hat{+} a_{2p}$ , hence by a positive induction  $a_2 \hat{+} a_1 \hat{+} a_4 \hat{+} a_3 \dots \hat{+} a_{2p} \hat{+} a_{2p-1}$ , hence by a negative induction  $a_2 \hat{+} a_1 \hat{-} a_4 \hat{+} a_3 \dots \hat{-} a_{2p} \hat{+} a_{2p-1}$ ; thus the other one of the two shapes is the shape of at least one tree in  $\Gamma(G)$ , and similarly if we start from the opposite one.  $\square$

## 6. CIRCULAR ORDER, AND DESCRIPTION OF THE GRAPH OF GRAPHS

**Lemma 6.1.** *Let  $G$  be a tree of relations,  $s, t, u$  its bijections; we say that  $b$  is the successor of  $a$  if  $b = tsu(a)$ ; this defines a total circular order on the vertices of  $G$ , invariant by any induction.*

**Proof**

We check first the invariance under the induction  $J_B$  for  $B$  a signed chain. Let  $s', t', u'$  be as in Lemma 3.7; suppose that  $a \in B$  and  $B$  is a negative chain; then  $u'a = sa$  is in  $B$ ,  $s'u'a = sussa = sua$  is in  $B$ , and thus  $t's'u'a = t'sua = tsua$ ; similarly, if  $a \in B$  and  $B$  is a positive chain, we get  $t's'u'a = sstsua = tsua$ ; if  $a$  is not in  $B$   $t's'u'a = tsua$ .

Because of this invariance and Proposition 5.2, we need only to check that we have a total circular order for trees without  $=$  edges. And for a tree  $a_1 \hat{-} a_2 \hat{+} a_3 \dots \hat{-} a_{2p}$  we get the order  $(a_1, a_3, \dots, a_{2p-1}, a_{2p}, a_{2p-2}, \dots, a_2, a_1)$ ; for a tree  $a_1 \hat{+} a_2 \hat{-} a_3 \dots \hat{+} a_{2p}$  we get the order  $(a_1, a_2, \dots, a_{2p-2}, a_{2p}, a_{2p-1}, \dots, a_3, a_1)$ ; for a tree  $a_1 \hat{-} a_2 \hat{+} a_3 \dots \hat{+} a_{2p+1}$  we get the order  $(a_1, a_3, \dots, a_{2p+1}, a_{2p}, a_{2p-2}, \dots, a_2, a_1)$ .  $\square$

The circular order described above may be described geometrically as follows: starting from any vertex  $x$  in a tree of relations  $G$ , we move from  $x$  to the vertex  $y$  where  $x$  and  $y$  are joined by a  $-$  edge. If  $x$  is not incident to a  $-$  edge, we take  $y$  to be  $x$ . Next we move from  $y$  to  $z$  where  $y$  and  $z$  are joined by a  $=$  edge. Again, if no such edge exists, we take  $z$  to be  $y$ . Finally we move from  $z$  to  $w$  where  $w$  is joined to  $z$  by a  $+$  edge. Then  $w$  is the successor of  $x$ . For the tree in Figure 1, the circular order is  $(1, 10, 3, 8, 5, 4, 6, 7, 9, 2, 1)$  while for every tree in Example 3.5 it is  $(1, 2, 3, 1)$ .

**Proposition 6.2.** *A tree of relations  $G$  is in  $\Gamma(G')$  if and only if the circular order of its vertices, given by  $t$ , is the same as the one of  $G'$ .*

**Proof**

By Lemma 6.1, the condition is necessary. By Proposition 5.2, it is enough to check the sufficiency for trees of relations without  $=$  edges.

If  $k$  is odd, the only shape is  $\hat{-}.\hat{+}....\hat{+}$ ; for it, every possible filling is of the form  $a_1\hat{-}a_2\hat{+}a_3... \hat{+}a_{2p+1}$  where the circular order  $(a_1, a_3, ..., a_{2p+1}, a_{2p}, a_{2p-2}, ..., a_2, a_1)$  coincides with the circular order on  $G'$  by assumption. This makes  $k$  trees of relations, one of which belongs to  $\Gamma(G')$  by Corollary 5.3. We call it  $a_1^1\hat{-}a_2^1\hat{+}a_3^1\hat{-}... \hat{+}a_{2p+1}^1$ ; all the others can be reached from it by induction, first by negative inductions going to  $a_1^1\hat{+}a_2^1\hat{-}a_3^1\hat{+}... \hat{+}a_{2p+1}^1$ , then by inducing successively  $1 \leq l \leq k$  times on the whole positive chain, and ending with negative inductions to replace  $=$  edges by  $-$  edges.

If  $k$  is even, take the shape  $\hat{-}.\hat{+}....\hat{-}$ ; for it, every possible filling is of the form  $a_1\hat{-}a_2\hat{+}a_3... \hat{-}a_{2p}$  where the circular order  $(a_1, a_3, ..., a_{2p-1}, a_{2p}, a_{2p-2}, ..., a_2, a_1)$  coincides with the order of  $G'$ ; this gives  $\frac{k}{2}$  trees of relations, one of them belongs to  $\Gamma(G')$  by Corollary 5.3, we call it  $a_1^1\hat{-}a_2^1\hat{+}a_3^1... \hat{-}a_{2p-1}^1$ ; all the others can be reached from it, first by negative inductions going to  $a_1^1\hat{+}a_2^1\hat{-}a_3^1\hat{+}... \hat{-}a_{2p-1}^1$ , then inducing successively  $2 \leq 2l \leq 2k$  times on the whole positive chain, and ending with negative inductions. The proof is similar for the shape  $\hat{+}.\hat{-}....\hat{+}$ .  $\square$

**Corollary 6.3.** *If  $G$  is a trees of relations with  $k$  vertices, then  $\Gamma(G)$  is obtained from  $\Gamma(k)$  by a renumbering of the vertices.  $G$  is in  $\Gamma(k)$  if and only if its circular order is  $(1, 2, ..., k, 1)$ .*

**Proof**

Immediate from Proposition 6.2 and computation of the circular order of  $G_0(k)$ .  $\square$

## 7. CARDINALITY OF THE GRAPH OF GRAPHS

**Lemma 7.1.** *Let  $F$  be a shape, and  $\sigma$  a bijection (vertex to vertex, edge to edge, preserving the labels) such that  $\sigma F = F$ . Then  $\sigma$  is an involution; if  $\sigma$  is not the identity,  $k$  is even, and  $F$  is the disjoint union of two subtrees  $F_1$  and  $F_2$  and an edge  $e$  such that  $\sigma e = e$ ,  $\sigma F_1 = F_2$ ,  $\sigma F_2 = F_1$ .*

*For a given  $F$  there is at most one such  $\sigma$  different from the identity.*

**Proof**

We consider  $\sigma$  as an isometry of a compact metric space by replacing edges by segments of length 1. As such,  $\sigma$  has a fixed point.

If the fixed point is a vertex,  $\sigma$  has to fix the adjacent edges (which are all distinct) hence the adjacent vertices, and, continuing this reasoning, we get that  $\sigma$  is the identity.

If the fixed point is an edge  $e$ , either  $\sigma$  fixes its two end vertices or it exchanges them. In the first case, again  $\sigma$  is the identity. In the second case, we get the desired decomposition given above and thus  $k$  is even. And  $\sigma^2$  fixes  $F_1$  and one of its vertices, thus is the identity on  $F_1$ , and similarly on  $F_2$ , thus on  $F$ .

For a given  $F$ , if it exists  $e$  is unique (as there is the same number of vertices on each side), and  $\sigma$  is defined uniquely on the adjacent edges, and thus everywhere.  $\square$

**Definition 7.2.** *A shape  $F$  is said to be symmetric if there exists a  $\sigma$  as in Lemma 7.1, different from the identity. We call such a  $\sigma$  a symmetry.*

**Lemma 7.3.** *If  $k$  is odd, every shape has  $k$  different fillings in  $\Gamma(k)$ ; if  $k$  is even, every symmetric shape has  $\frac{k}{2}$  different fillings and every non-symmetric shape has  $k$  different fillings.*



**Proof**

Given a shape  $F$ , we fix a way to relate it by a sequence of inductions, as in the proof of Proposition 5.2, to a shape  $F'$  without  $=$  edges; if  $F$  is non-symmetric, this defines a map  $\phi$  from the fillings  $G$  of  $F$  to the fillings  $G'$  of  $F'$  (if  $\phi$  gives a different image to one filling written in two different ways, this defines a symmetry on  $F$ ).

If  $F$  and  $F'$  are non-symmetric, then  $\phi$  is a bijection. This happens if  $k$  is odd by Lemma 7.1, and in this case we need only consider the shape  $F' = \hat{-}.\hat{+}....\hat{+}$ ; we have seen in the proof of Proposition 6.2 that it has  $k$  fillings, thus so has  $F$ .

If  $k$  is even, and  $F$  is non-symmetric; then  $F'$  is one of the (symmetric) shapes  $\hat{-}.\hat{+}....\hat{-}$  and  $\hat{+}.\hat{-}....\hat{+}$ , each of which has  $\frac{k}{2}$  fillings, and  $\phi$  is two-to-one, hence  $F$  has  $k$  fillings.

If  $k$  is even, and  $F$  is symmetric; then we can relate  $F$  to an  $F'$  without  $=$  edges such that at each stage the shape is symmetric (by working simultaneously on one extremal  $\hat{=}$  in  $F_1$  and its image under the symmetry); thus our  $\phi$  sends the symmetry of  $F$  to the (unique) symmetry of  $F'$ , thus  $\phi$  is well defined and one-to-one, hence  $F$  has  $\frac{k}{2}$  fillings.  $\square$

**Corollary 7.4.**  $\Gamma(k)$  contains  $k$  trees of relations without  $=$  edges.

**Proposition 7.5.**

$$\#\Gamma(k) = \text{Cat}_{k+1} - \text{Cat}_k$$

where  $\text{Cat}_k = \frac{(2k)!}{(k+1)(k!)^2}$  is the  $k$ -th Catalan number.

The number of different shapes is  $\frac{\text{Cat}_{k+1}-\text{Cat}_k}{k}$  if  $k$  is odd, and  $\frac{\text{Cat}_{k+1}-\text{Cat}_k}{k} + \frac{3}{2}\text{Cat}_{\frac{k}{2}}$  if  $k$  is even.

**Proof**

We count first the number  $\rho(k)$  of rooted shapes (where we have specified one vertex, see Definition 5.1 above) on  $k$  letters. We define four quantities,

- $\rho_0(k)$  is the number of rooted shapes with no edge from the root;
- $\rho_1(k)$  is the number of rooted shapes with one edge from the root, labelled  $+$ , resp.  $=$ , resp.  $-$  (these three numbers being equal);
- $\rho_2(k)$  is the number of rooted shapes with two edges from the root, labelled  $+$  and  $-$ , resp.  $=$  and  $-$ , resp.  $+$  and  $=$ ;
- $\rho_3(k)$  is the number of rooted shapes with three edges from the root;
- $\zeta(k)$  is the number of rooted shapes with no edge from the root labelled  $+$ , resp.  $=$ , resp.  $-$ .

We have

- $\rho(k) = \rho_0(k) + 3\rho_1(k) + 3\rho_2(k) + \rho_3(k)$ ;
- $\zeta(k) = \rho_0(k) + 2\rho_1(k) + \rho_2(k)$ ;
- $\rho_0(1) = 1, \rho_0(k) = 0$  if  $k > 1$ ;
- $\rho_1(k) = \zeta(k-1)$ ;
- $\rho_2(k) = \sum_{p+q=k-1} \zeta(p)\zeta(q)$ ;
- $\rho_3(k) = \sum_{p+q+r=k-1} \zeta(p)\zeta(q)\zeta(r)$ .

In terms of power series  $R_i = \sum_{k \geq 1} \rho_i(k)X^k$ ,  $Z = \sum_{k \geq 1} \zeta(k)X^k$  we get  $R_1 = XZ$ ,  $R_2 = XZ^2$ ,  $R_3 = XZ^3$ . Hence  $Z = X(Z^2 + 2Z + 1)$ ,  $XZ^2 + (2X - 1)Z + X = 0$ , and the only admissible solution is  $Z = \frac{1-\sqrt{1-4X}}{2X} - 1$ . This gives  $\zeta(k) = \text{Cat}_k$ .

Now  $R = X(Z+1)^3 = Z(Z+1) = \frac{Z}{X} - Z - 1$ , thus  $\rho(k) = \zeta(k+1) - \zeta(k)$ .

If  $k$  is odd, one shape gives  $k$  rooted shapes, so the number of shapes is  $\frac{\rho(k)}{k}$  and the number of trees of relations in  $\Gamma_k$  is  $\rho(k)$  by Lemma 7.3.

If  $k$  is even, a symmetric shape is made with one central edge (3 choices) and a compatible rooted shape on  $\frac{k}{2}$  letters, thus the number  $\tau(k)$  of symmetric shapes is  $3\zeta(\frac{k}{2})$ . A symmetric shape corresponds to  $\frac{k}{2}$  rooted shapes and a non-symmetric shape corresponds to  $k$  rooted shapes; if  $\psi(k)$  is the number of non-symmetric shapes,  $\rho(k) = \frac{k}{2}\tau(k) + k\psi(k)$ . By Lemma 7.3 the number of trees of relations in  $\Gamma_k$  is thus  $\rho(k)$ , and, by computing  $\psi(k)$ , we get the claimed number of shapes.  $\square$

Note that if we fix an order on the labels, such as  $-$  smaller than  $=$  and  $=$  smaller than  $+$ , the rooted shapes whose root has no  $-$  (resp.  $=$ ,  $+$ ) edge, whose number is  $\zeta(k)$ , are in bijection with the set of ordered (incomplete) binary trees (suspending them by the root) and this proves again that  $\zeta(k)$  is the  $k$ -th Catalan number; see [9] for example.

The sequence of numbers of shapes begins by 1, 3, 3, 10, 18, 57... and seems to be a new sequence, not yet in the *On-line Encyclopedia of Integer Sequences* [6].

The third author together with R. Marsh and S. Schroll have shown that the set of shapes on  $k$ -vertices is in bijection with the set of all labelled triangulations of a regular  $(k+2)$ -sided polygon up to rotations, where the edges of each triangle are labelled  $+$ ,  $=$ ,  $-$  in the clockwise direction. Although we do not exploit this point in this paper, the induction mapping may be defined on shapes by simply forgetting the labeling of the vertices. In this way, the induction mapping may be regarded as a mapping defined on labelled triangulated polygons. This alternative perspective leads to another proof of the results obtained in §7 by way of a formula due to Brown [1] for the number of triangulations of a regular  $(k+2)$ -sided polygon up to rotations which incidentally corresponds with the number of isomorphism classes of (basic) cluster-tilted algebras of type  $A_{k-1}$  [8].

## 8. SECONDARY STRUCTURES OF GENETIC SEQUENCES

In §6 we associated a circular order to a tree structure; now we start from a circular order and a structure imitated from the secondary structures of RNA (see for example [2]) and get trees of relations. Namely

**Definition 8.1.** *Let  $S$  be the periodic circular string on three symbols  $(XYZ)^k$ ; we equip it with an origin and denote it by  $X_1Y_1Z_1\dots X_kY_kZ_k$ . A pseudo knot-free secondary structure  $\Sigma$  on  $S$  is a set of links between two different instances of symbol  $X$ , or two different instances of symbol  $Y$ , or two different instances of symbol  $Z$ , such that any pair of distinct links, drawn inside the circle, have an empty intersection.*

**Proposition 8.2.** *Let  $\Sigma$  be a pseudo knot-free secondary structure as above. We define a graph  $G$  by putting  $a\hat{+}b$ , resp.  $a\hat{=}b$ , resp.  $a\hat{-}b$ , if there is a link between  $X_a$  and  $X_b$ , resp.  $Y_a$  and  $Y_b$ , resp.  $Z_a$  and  $Z_b$ . Then  $G$  is a disjoint union of trees of relations  $G_i$ ,  $1 \leq i \leq d$ , such that if  $i \neq j$ ,  $a_1$  and  $a_2$  are vertices of  $G_i$ , and  $b_1$  and  $b_2$  are vertices of  $G_j$ , we cannot have  $a_1 < b_1 < a_2 < b_2$  in the circular order  $(1, \dots, k, 1)$ . We can then define the maps  $s$ ,  $t$ ,  $u$  as in §3. For any vertex  $a$  of  $G_i$ ,  $tsu(a)$  is the next element of  $G_i$  in the circular order  $(1, \dots, k, 1)$ .*

### Proof

Note first that two adjacent edges in  $G$  have different labels, otherwise two different links in  $\Sigma$  have a nonempty intersection.

We show that  $G$  has no loop; suppose there is a loop with vertices  $a_1, \dots, a_r$ ; then it cannot be the case that all of its edges are labelled  $=$ . If there is an edge  $a_i \hat{+} b_i$ , there is a link between  $X_{a_i}$  and  $X_{b_i}$ . Then any link from  $Y_{a_i}$  or  $Z_{a_i}$  not intersecting this  $X$ -link goes to  $Y_c$  or  $Z_c$  for some  $a_i < c < b_i$  (in the circular order), while any link from  $Y_{b_i}$  or  $Z_{b_i}$  not intersecting the  $X$ -link goes to  $Y_d$  or  $Z_d$  for some  $b_i < d < a_i$  (in the circular order), and there is no way to close the loop. A similar reasoning applies if there is an edge  $a_i \hat{-} b_i$ .

Thus  $G$  is indeed a disjoint union of trees of relations  $G_i$ , and the condition on four vertices is necessary to avoid nonempty intersections. For a vertex  $a$  of  $G_i$ , either  $u(a) = a$  or there is a link between  $Z_a$  and  $Z_{u(a)}$ , either  $su(a) = u(a)$  or there is a link between  $Y_{su(a)}$  and  $Y_{u(a)}$ , either  $tsu(a) = su(a)$  or there is a link between  $Z_{su(a)}$  and  $Z_{tsu(a)}$ . Thus  $tsu(a)$  is a vertex of  $G_i$ , and, to avoid nonempty intersections, every vertex situated strictly between  $a$  and  $tsu(a)$  can be linked only to another vertex strictly between  $a$  and  $tsu(a)$ , thus such a vertex is not in  $G_i$ .  $\square$

**Proposition 8.3.** *Let  $G$  be a disjoint union of trees of relations  $G_i$ ,  $1 \leq i \leq d$ ; we equip the set of all vertices of the  $G_i$  with any circular order compatible with the circular order  $tsu$  defined on each  $G_i$ , and such that if  $a_1$  and  $a_2$  are vertices of  $G_i$ ,  $b_1$  and  $b_2$  vertices of  $G_j$ , we do not have  $a_1 < b_1 < a_2 < b_2$  in this order. We define a link between  $X_a$  and  $X_b$ , resp.  $Y_a$  and  $Y_b$ , resp.  $Z_a$  and  $Z_b$ , whenever there is an edge  $a \hat{+} b$ , resp.  $a \hat{=} b$ , resp.  $a \hat{-} b$  in  $G$ ; then we get a pseudo knot-free secondary structure as above.*

#### Proof

What we have to prove is that any two distinct links have a nonempty intersection. When we have one tree of relations  $G$ , this is trivially true if  $G$  has no  $=$  edge; and the definition of the induction in §3 implies that this property is stable under induction, thus the result follows by Proposition 5.2. When we have several trees, the condition on the order allows to mix the structures without creating intersections.  $\square$

**Proposition 8.4.** *A single tree of relations defines a pseudo knot-free secondary structure which is maximal: no link can be added on the same set of vertices.*

#### Proof

Any extra link would add an extra edge but by Proposition 8.2 the new graph has to be a union of trees.  $\square$

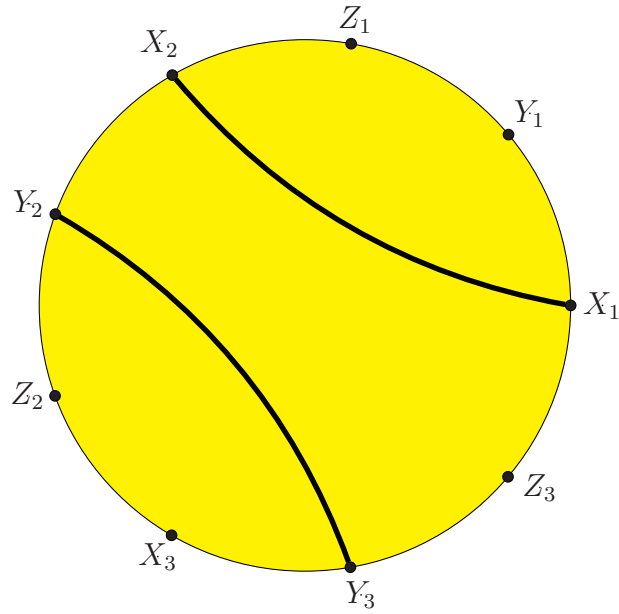
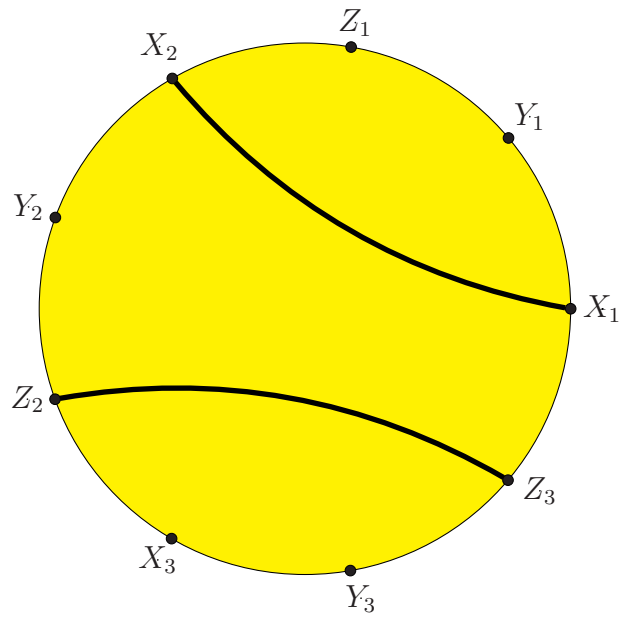
Figures 18 and 19 illustrate pseudo knot-free structures corresponding to the trees of relations  $1 \hat{+} 2 \hat{=} 3$  and  $1 \hat{+} 2 \hat{-} 3$  in  $\Gamma(3)$ . Figure 20 illustrates an intersection between the arcs for the tree of relations  $3 \hat{-} 1 \hat{+} 2$  not contained in  $\Gamma(3)$  (of course there would be no intersection if we used the circular order defined by  $3 \hat{-} 1 \hat{+} 2$ , namely  $(1, 3, 2, 1)$ ).

**Example 8.5.** The forest  $1 \hat{+} 2 \hat{-} 3, 4 \hat{=} 5$  also defines a maximal pseudo knot-free secondary structure.

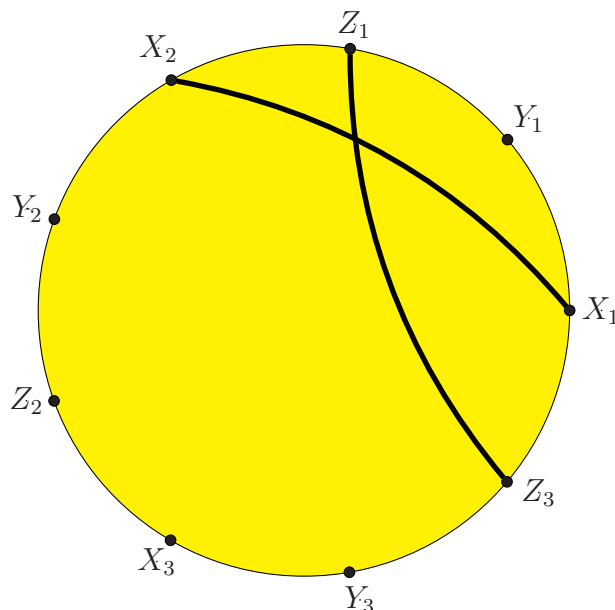
Thus, to ensure that a given structure  $\Sigma$  on  $(XYZ)^k$  corresponds to a single tree of relations, we need to specify that  $\Sigma$  has  $k - 1$  links.

## REFERENCES

- [1] G.W. BROWN, Enumerations of triangulations of the disk, *Proc. London Math. Soc.* (3) **14** (1964), p. 746–768.

FIGURE 18.  $1\hat{+}2\hat{=}3$ .FIGURE 19.  $1\hat{+}2\hat{-}3$ .

- [2] A. CONDON, Problems on RNA secondary structure prediction and design, *Automata, languages and programming*, Lecture Notes in Comput. Sci. **2719** (2003), p. 22–32.
- [3] S. FERENCZI, L.Q. ZAMBONI, Structure of  $k$ -interval exchange transformations: induction, trajectories, and distance theorems, *J. Anal. Math.* **112** (2010), p. 289–328.
- [4] S. FERENCZI, L.Q. ZAMBONI, Eigenvalues and simplicity of interval exchange transformations, to appear in *Ann. Sci. Ec. Norm. Sup. (4)* **44** (2011).

FIGURE 20.  $3\hat{-}1\hat{+}2$ .

- [5] R. MARSH, S. SCHROLL, RNA secondary structures, polygon dissections and clusters, preprint arXiv:1010.3763v1 [math.CO], 2010.
- [6] N.J.A. SLOANE, The On-Line Encyclopedia of Integer Sequences, (see <http://www.research.att.com/njas/sequences/>).
- [7] R.P. STANLEY, Enumerative combinatorics. Vol. 2, with a foreword by Gian-Carlo Rota and appendix 1 by Sergey Fomin. Cambridge Studies in Advanced Mathematics, 62. Cambridge University Press, Cambridge, 1999.
- [8] H.A. TORKILDSEN, Counting cluster-tilted algebras of type  $A_n$ , *Int. Electron. J. Algebra* **4** (2008), p. 149–158.
- [9] B. VANCE, Counting ordered trees by permuting their parts, *Amer. Math. Monthly* **113** (2006), p. 329–335.

INSTITUT DE MATHÉMATIQUES DE LUMINY, CNRS UMR 6206, CASE 907, 163 AV. DE LUMINY, F13288 MARSEILLE CEDEX 9 (FRANCE) AND FÉDÉRATION DE RECHERCHE DES UNITÉS DE MATHÉMATIQUES DE MARSEILLE, CNRS FR 2291

*E-mail address:* `cassaigne@iml.univ-mrs.fr`

INSTITUT DE MATHÉMATIQUES DE LUMINY, CNRS UMR 6206, CASE 907, 163 AV. DE LUMINY, F13288 MARSEILLE CEDEX 9, FRANCE AND FÉDÉRATION DE RECHERCHE DES UNITÉS DE MATHÉMATIQUES DE MARSEILLE, CNRS FR 2291

*E-mail address:* `ferenczi@iml.univ-mrs.fr`

UNIVERSITÉ DE LYON, UNIVERSITÉ LYON 1, CNRS UMR 5208, INSTITUT CAMILLE JORDAN, 43 BOULEVARD DU 11 NOVEMBRE 1918, F69622 VILLEURBANNE CEDEX, FRANCE. ALSO, DEPARTMENT OF MATHEMATICS, FUNDIM, UNIVERSITY OF TURKU, FIN-20014 TURKU, FINLAND.

*E-mail address:* `zamboni@math.univ-lyon1.fr`, `luca.zamboni@utu.fi`